

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: HighTech Receptro AB
- (B) STREET: c/o Active, Skeppsbron 2
- (C) CITY: MALMO
- (E) COUNTRY: SWEDEN
- (F) POSTAL CODE (ZIP): 211 20
- (G) TELEPHONE: 040/35 07 00
- (H) TELEFAX: 040/ 23 74 05
- (I) TELEX: 32637 Active S

(ii) TITLE OF INVENTION: Hybridprotein

(iii) NUMBER OF SEQUENCES: 1

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: SE PCT/SE93/00375

(vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: SE 9201331-7
- (B) FILING DATE: 28-APR-1992

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Ala	Val	Glu	Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser
1				5					10					15	
Glu	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser
			20					25						30	
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu
			35				40						45		

Ala	Tyr	Ala	Tyr	Ala	Asp	Thr	Leu	Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr		
50						55					60						
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly		
65					70					75					80		
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala		
				85					90					95			
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly		
			100					105					110				
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu		
		115					120					125					
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr		
	130					135					140						
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro		
145					150					155					160		
Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys		
				165					170					175			
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu		
			180					185					190				
Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr		
		195					200					205					
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly		
	210					215					220						
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala		
225					230					235					240		
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly		
				245					250					255			
Thr	Phe	Ala	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu		
			260					265					270				
Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr	Ala	Asp	Leu	Glu	Asp	Gly	Gly	Tyr		
		275					280					285					
Thr	Ile	Asn	Ile	Arg	Phe	Ala	Gly	Lys	Lys	Val	Asp	Glu	Lys	Pro	Glu		
	290					295					300						
Glu																	
305																	

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 921 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Escherichia coli* LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA	50
AGAAGAAGTA ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC	100
AAACTGCAGA ATTCAAAGGA ACATTTGAAA AAGCAACATC AGAAGCTTAT	150
GCGTATGCAG ATACTTTGAA GAAAGACAAT GGAGAATATA CTGTAGATGT	200
TGCAGATAAA GGTATACTT TAAATATTAA ATTTGCTGGA AAAGAAAAAA	250
CACCAGAAGA ACCAAAAGAA GAAGTTACTA TTAAAGCAAA CTTAATCTAT	300
GCAGATGGAA AAACACAAAC AGCAGAATTC AAAGGAACAT TTGAAGAAGC	350
AACAGCAGAA GCATACAGAT ATGCAGATGC ATTAAAGAAG GACAATGGAG	400
AATATACAGT AGACGTTGCA GATAAAGGTT ATACTTTAAA TATTAAATTT	450
GCTGGAAAAG AAAAAACACC AGAAGAACCA AAAGAAGAAG TTACTATTAA	500
AGCAAACCTTA ATCTATGCAG ATGGAAAAAC ACAAACAGCA GAATTCAAAG	550
GAACATTTGA AGAAGCAACA GCAGAAGCAT ACAGATATGC TGACTTATTA	600
GCAAAAGAAA ATGGTAAATA TACAGTAGAC GTTGCAGATA AAGGTTATAC	650
TTTAAATATT AAATTTGCTG GAAAAGAAAA AACACCAGAA GAACCAAAAG	700
AAGAAGTTAC TATTAAAGCA AACTTAATCT ATGCAGATGG AAAA ACTCAA	750
ACAGCAGAGT TCAAAGGAAC ATTTGCAGAA GCAACAGCAG AAGCATACAG	800
ATACGCTGAC TTATTAGCAA AAGAAAATGG TAAATATACA GCAGACTTAG	850
AAGATGGTGG ATACACTATT AATATTAGAT TTGCAGGTAA GAAAGTTGAC	900
GAAAAACCAG AAGAATAATA A	921

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Escherichia coli* LE392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala	Val	Glu	Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser	1	5	10	15
Glu	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser	20	25	30	
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu	35	40	45	
Ala	Tyr	Ala	Tyr	Ala	Asp	Thr	Leu	Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	50	55	60	
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	65	70	75	80
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	85	90	95	
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	100	105	110	
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu	115	120	125	
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	130	135	140	
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	145	150	155	160
Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	165	170	175	
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	180	185	190	
Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr	195	200	205	
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	210	215	220	
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	225	230	235	240
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	245	250	255	
Thr	Phe	Ala	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	260	265	270	

Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
 275 280 285

Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu
 290 295 300

Glu Pro Met Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys
 305 310 315 320

Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val
 325 330 335

Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr
 340 345 350

Asp Asp Ala Thr Lys Thr Phe Thr Val Thr Glu Lys Pro Glu Val Ile
 355 360 365

Asp Ala Ser Glu Leu Thr Pro Ala Val Thr Thr Tyr Lys Leu Val Ile
 370 375 380

Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr Thr Lys Ala Val Asp Ala
 385 390 395 400

Glu Thr Ala Glu Lys Ala Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val
 405 410 415

Asp Gly Val Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val Thr
 420 425 430

Glu Met

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli L392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA	50
AGAAGAAGTA ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC	100
AAACTGCAGA ATTCAAAGGA ACATTTGAAA AAGCAACATC AGAAGCTTAT	150
GCGTATGCAG ATACTTTGAA GAAAGACAAT GGAGAATATA CTGTAGATGT	200

TGCAGATAAA	GGTTATACTT	TAAATATTAA	ATTTGCTGGA	AAAGAAAAAA	250
CACCAGAAGA	ACCAAAAGAA	GAAGTTACTA	TTAAAGCAAA	CTTAATCTAT	300
GCAGATGGAA	AAACACAAAC	AGCAGAATTC	AAAGGAACAT	TTGAAGAAGC	350
AACAGCAGAA	GCATACAGAT	ATGCAGATGC	ATTAAAGAAG	GACAATGGAG	400
AATATACAGT	AGACGTTGCA	GATAAAGGTT	ATACTTTAAA	TATTAAATTT	450
GCTGGAAAAG	AAAAAACACC	AGAAGAACCA	AAAGAAGAAG	TTACTATTAA	500
AGCAAACCTTA	ATCTATGCAG	ATGGAAAAAC	ACAAACAGCA	GAATTCAAAG	550
GAACATTTGA	AGAAGCAACA	GCAGAAGCAT	ACAGATATGC	TGACTTATTA	600
GCAAAAGAAA	ATGGTAAATA	TACAGTAGAC	GTTGCAGATA	AAGGTTATAC	650
TTTAAATATT	AAATTTGCTG	GAAAAGAAAA	AACACCAGAA	GAACCAAAAG	700
AAGAAGTTAC	TATTAAAGCA	AACTTAATCT	ATGCAGATGG	AAAAACTCAA	750
ACAGCAGAGT	TCAAAGGAAC	ATTTGCAGAA	GCAACAGCAG	AAGCATACAG	800
ATACGCTGAC	TTATTAGCAA	AAGAAAATGG	TAAATATACA	GCAGACTTAG	850
AAGATGGTGG	ATACACTATT	AATATTAGAT	TTGCAGGTAA	GAAAGTTGAC	900
GAAAACCAG	AAGAACCCAT	GGACACTTAC	AAATTAATCC	TTAATGGTAA	950
AACATTGAAA	GGCGAAACAA	CTACTGAAGC	TGTTGATGCT	GCTACTGCAG	1000
AAAAAGTCTT	CAAACAATAC	GCTAACGACA	ACGGTGTTGA	CGGTGAATGG	1050
ACTTACGACG	ATGCGACTAA	GACCTTTACA	GTTACTGAAA	AACCAGAAGT	1100
GATCGATGCG	TCTGAATTAA	CACCAGCCGT	GACAACTTAC	AAACTTGTTA	1150
TTAATGGTAA	AACATTGAAA	GGCGAAACAA	CTACTAAAGC	AGTAGACGCA	1200
GAAACTGCAG	AAAAAGCCTT	CAAACAATAC	GCTAACGACA	ACGGTGTTGA	1250
TGGTGTTTGG	ACTTATGATG	ATGCGACTAA	GACCTTTACG	GTAAGTAAA	1300
TGTAATAA					1308